

Witold MAŃCZAK (Kraków)

**DÉVELOPPEMENT PHONÉTIQUE IRRÉGULIER  
DÛ À LA FRÉQUENCE  
ET DICTIONNAIRES ÉTYMOLOGIQUES**

Vers la fin des années cinquante, nous sommes arrivé à la conclusion que, dans toutes les langues, le développement phonétique irrégulier dû à la fréquence est, à côté du développement phonétique régulier et du développement analogique, le troisième facteur essentiel qui décide de la forme des mots. Nous y avons consacré trois monographies (Mańczak 1969, 1977 et 1987) et de nombreux articles. Il y a six arguments à l'appui de notre opinion.

1° S'il existe, pour la langue et la période en question, un dictionnaire de fréquence, on peut en user parce que la grande majorité des mots subissant un développement phonétique irrégulier dû à la fréquence se trouve parmi les mots les plus employés. Par exemple, en français moderne, ces mots se présentent comme suit :

1 <sup>er</sup> mille	99	86%
2 <sup>e</sup> mille	9	8%
3 <sup>e</sup> mille	4	3%
4 <sup>e</sup> mille	2	2%
5 <sup>e</sup> mille	1	1%
6 <sup>e</sup> mille	—	

Si, par hasard, quelqu'un doutait que ces données statistiques prouvent qu'il y a un lien entre les réductions irrégulières et la fréquence des mots, on pourrait attirer son attention sur le fait que l'application du test  $\chi^2$  donne comme résultat  $\chi^2 = 409,55 > 11,07$  (5 degrés de liberté, probabilité 0,05).

2° Si le morphème, mot ou groupe de mots apparaît dans une langue donnée sous une double forme, régulière et irrégulière, le développement phonétique irrégulier se caractérise par le fait que la forme régulière est en général plus employée que la forme normale. Par exemple, *Franç-ais* est plus employé

que *Franç-ois*, *se* est plus employé que *soi*, et *monsieur* est plus utilisé que *monseigneur*.

3° Si les réductions irrégulières ont lieu à l'intérieur d'un paradigme ou d'une famille de mots, les abrègements irréguliers se produisent plus souvent dans des formes plus fréquentes que dans des formes moins employées. Prenons en considération le développement du prés. de l'ind. *habeo*, *habes*, etc. et du prés. du subj. *habeam*, *habeas*, etc. Parmi les formes italiennes *ho*, *hai*, *ha*, *abbiamo*, *avete*, *hanno* sont irréguliers *ho*, *hai*, *ha*, *hanno*, ce qui s'explique par le fait que le singulier est plus employé que le pluriel et la 3<sup>e</sup> personne est plus employée que les autres. Les formes du prés. du subj. *abbia*, *abbia*, *abbia*, *abbiamo*, *abbiate*, *abbiano* sont toutes régulières, ce que s'explique par le fait que le subjonctif est moins employé que l'indicatif. En français, le groupe intervocalique *-tr-* aboutit normalement à *-rr-*, cf. *petram* > *pierre*. Il en résulte que *père* < *patrem* présente une réduction, alors que le développement de *parrain* < \**patrinum* est normal. Il est évident que l'irrégulier *père* est plus fréquemment usité que le régulier *parrain*.

4° Si, pour une langue donnée, on dispose à la fois d'un dictionnaire de fréquence et d'un dictionnaire inverse, il est instructif d'examiner des séries de mots commençant par la même lettre ou les mêmes lettres ou bien des séries de mots terminés par la même lettre ou les mêmes lettres. Grâce au dictionnaire de Thorndike et Lorge (1944) et celui de Lehnert (1971), il nous a été possible d'établir ce qui suit :

En anglais, il y a 16 verbes faibles monosyllabiques terminés en *-ay*, par exemple *play*, parmi lesquels il n'y en a qu'un seul dont les formes présentent des réductions irrégulières, à savoir *say* (*says*, *said*), et *say* est le verbe le plus employé en *-ay*.

Il y a 5 mots terminés en *-ayer*, par exemple *layer*, parmi lesquels il n'y en a qu'un seul dont la prononciation soit réduite, à savoir *prayer*, et *prayer* est le mot le plus utilisé parmi les mots en *-ayer*.

Il y a 9 mots en *-een*, par exemple *seen*, parmi lesquels il n'y en a qu'un seul dont la prononciation puisse être réduite, à savoir *been*, et *been* est le mot le plus employé parmi les mots en *-een*.

Dans le dictionnaire de Lehnert, il y a plus de 200 féminins en *-ess*, par exemple *princess*, parmi lesquels il n'y en a que deux dont la prononciation est réduite, à savoir *mistress* > *Mrs.* et *Miss*, et *Mrs.* ainsi que *Miss* sont les féminins le plus fréquemment usités.

Il y a 34 mots monosyllabiques en *-f*, par exemple *if*, parmi lesquels il n'y en a qu'un seul dont la consonne finale ait subi une sonorisation irrégulière, à savoir *of*, et *of* est le mot le plus employé parmi les mots en *-f*.

Il y a 34 mots en *-ill*, par exemple *still*, parmi lesquels il n'y en a qu'un seul dont la prononciation puisse subir une réduction, à savoir *will*, et *will* est le mot le plus employé en *-ill*.

Il existe 15 mots terminés par une consonne fricative + *-in*, par exemple *coffin*, parmi lesquels il n'y en a que deux qui aient subi une réduction, à savoir *cousin* et *basin*, et, du point de vue de la fréquence, *cousin* et *basin* occupent, parmi ces mots, la première et la deuxième position.

Il y a 12 mots monosyllabiques en *-ine*, par exemple *line*, parmi lesquels il n'y en a qu'un seul qui ait subi une réduction, à savoir *mine* > *my*, et *my* est plus utilisé que tous les mots en *-ine*.

Dans le dictionnaire inverse de Lehnert, il y a presque 700 dérivés en *-iness*, par exemple *happiness*, parmi lesquels il n'y en a qu'un seul qui ait subi une réduction, à savoir *business*, et *business* est le mot le plus fréquemment usité parmi les mots en *-iness*.

Il y a 36 mots en *-ire*, par exemple *fire*, parmi lesquels il n'y en a qu'un seul qui ait subi un abrègement, à savoir *sire* > *sir*, et, du point de vue de la fréquence, *sir* occupe, parmi ces mots, la deuxième position.

Il y a 11 mots monosyllabiques en *-one*, par exemple *tone*, parmi lesquels il n'y en a qu'un seul dont la voyelle ait été abrégée, à savoir *gone*, et *gone* est le mot le plus employé parmi les mots en *-one*.

Il y a 9 mots monosyllabiques en *-our*, parmi lesquels il n'y en a qu'un seul où *-our* puisse se réduire en [ə], à savoir *your*, et *your* est le mot le plus utilisé parmi les mots en *-our*.

Il y a 6 mots en *-over*, par exemple *clover*, parmi lesquels il n'y en a qu'un seul dont la prononciation puisse être réduite, à savoir *over* > *o'er*, et *over* est le mot le plus employé parmi ces mots.

Il y a 7 mots en *-teen*, par exemple *thirteen*, mais le numéral *ten*, qui, au point de vue étymologique, est la même chose que *-teen*, a subi un abrègement, et *ten* est plus employé que les numéraux en *-teen*.

Il y a 35 mots en *-ther*, par exemple *father*, parmi lesquels il n'y en a qu'un seul, à savoir a. angl. *wither*, qui ait été réduit en *with*, et *with* est plus employé que les mots en *-ther*.

Il y a 6 composés qui présentent un *al-* comme leur premier terme, par exemple *already*. Parmi ces composés, il n'y en a qu'un seul où la voyelle de *al-* se réduise en [ə], à savoir *alone*, et *alone* est le plus employé parmi les composés en *al-*.

Il y a 34 prétérits commençant par *be-*, par exemple *believed*, parmi lesquels il n'y en a qu'un seul dont le préfixe ait été, dans le passé, quelquefois supprimé, à savoir *began* > *gan*, et *began* est la forme la plus employée parmi celles qui commencent par *be-*.

Il y a 15 mots commençant par *count-*, par exemple *county*, parmi lesquels il n'y en a qu'un seul qui ait subi une réduction, à savoir *country*, et *country* est le plus employé parmi les mots commençant par *count-*.

Il y a plus de 1000 mots d'origine germanique commençant par *h-*, parmi lesquels il n'y en a qu'un seul qui, dans l'orthographe officielle, ait perdu *h*, à savoir *it* < *hit*, et *it* est plus employé que tous les mots où *h-* a persisté.

Il y a 18 mots commençant par *spr-*, par exemple *spring*. Parmi ces mots, il y en a eu, en a. anglais, aussi le verbe *sprecan*, dont la partie initiale a subi, par la suite, une réduction, comme en témoigne la forme moderne *speak*, et *speak* est plus employé que tous les mots où *spr-* s'est maintenu jusqu'à nos jours.

Il y a 62 mots commençant par *sw-*, par exemple *sweet*. En a. anglais, ces mots ont été un peu plus nombreux parce que les mots modernes *so*, *such*, *sister* et *sword* présentaient *sw-*. Parmi quelques dizaines de mots qui commencent de nos jours ou ont commencé dans le passé par *sw-*, du point de vue de la fréquence, *so* occupe la première place, *such*, la deuxième, *sister*, la quatrième et *sword*, la dix-neuvième, à ceci près qu'il est indubitable qu'autrefois *sword* a été plus employé que maintenant.

Parmi les mots commençant par *th-*, la plupart a conservé la prononciation régulière sourde, par exemple *think*, tandis que la minorité présente une sonorisation anormale, par exemple *that*. En prenant en considération tous les mots commençant par *th-*, nous avons établi que la fréquence moyenne des mots ayant subi une sonorisation irrégulière est d'environ 9000, tandis que la fréquence moyenne des mots qui ont conservé la prononciation régulière sourde est d'environ 300.

Il y a 13 mots dérivés du numéral *two*, par exemple *twelve*, *twice*, *twilight*, etc. Dans tous les dérivés le groupe *tw-* s'est maintenu tel quel, tandis que, dans le numéral *two*, le groupe initial *tw-* a subi une réduction, et *two* est plus utilisé que tous les dérivés en question.

Évidemment, il est impossible d'imaginer que toutes ces données statistiques que nous venons de citer soient dues au hasard. Au contraire, elles témoignent de ce que, entre la fréquence d'emploi et les réductions irrégulières, il existe un lien de cause à effet.

5° A côté du développement phonétique irrégulier dû à la fréquence, il existe une autre évolution irrégulière, consistant en des accidents phonétiques connus depuis longtemps sous le nom d'assimilations, dissimilations ou métathèses, en des formes hypercorrectes ou expressives. Toute cette évolution se caractérise par le fait qu'elle a lieu, dans des langues diverses, dans les mots les plus divers. *Chercher* < *cercher* présente une assimilation, *faible* < *flebilem* une dissimilation, *troubler* < *\*turbulare* une métathèse, tandis que *s* dans *besicles* est hypercorrect et *h* dans *herse* est censé être expressif. Mais il serait difficile

de trouver, dans une autre langue indo-européenne, un mot signifiant « faible » avec une dissimilation, un mot signifiant « troubler » avec une métathèse ou un mot signifiant « herse » avec un phonème d'origine expressif. Bref, il n'y a aucun parallélisme entre les irrégularités dites assimilations, dissimilations, métathèses, etc. En revanche, le développement phonétique irrégulier dû à la fréquence a lieu, dans des langues diverses, d'une manière plus ou moins parallèle parce que, malgré les différences qui séparent des communautés linguistiques, les mots les plus fréquents sont partout plus ou moins les mêmes. Par exemple, le verbe signifiant « parler » présente dans beaucoup de langues des réductions irrégulières, cf. fr. *parler*, it. *parlare* < *parabolare*, des formes de *narrare* devenu en sarde *nàrrere* (*nau*, *nas*, *nat*, etc.), lat. *ajo* < *\*agiō* (en face du régulier *adagium*), angl. *says*, *said* (en regard du régulier *lays*, *laid*, où la diphtongue a persisté), russe dial. *gyt* < *gryt* < *govorit*, a. polonais *pry* < *prawi*, etc.

6° De deux formes à développement phonétique irrégulier dû à la fréquence, la forme plus fréquemment usitée existe, en principe, sur un territoire plus vaste que la forme moins employée. Dans un texte français, nous avons trouvé 101 infinitifs en *-er*, 36 en *-ir* et 18 en *-oir*. En ce qui concerne les parlers français, on sait qu'il y en a où le suffixe de l'infinitif subit la chute irrégulière du *-r* non seulement dans les infinitifs du type *aller*, mais aussi dans ceux du type *dormir* et sporadiquement même dans ceux du type *avoir*. Dans l'*ALF*, nous avons trouvé 291 points où coexistent les formes d'infinitifs avec ou sans *-r*, à ceci près que la répartition des points où *-r* se maintient se présente de la façon suivante :

<i>aller</i>	—
<i>dormir</i>	103
<i>avoir</i>	283

Il en résulte que l'aire de la prononciation *alle(r)* est plus grande que celle de la prononciation *dormi(r)*, et celle-ci est plus grande que celle de la prononciation *avoi(r)*. Il existe donc un lien entre la fréquence des types d'infinitifs *aller*, *dormir*, *avoir* et la grandeur des aires où les suffixes de ces infinitifs subissent la chute irrégulière de *-r*.

Il y a déjà longtemps, nous avons publié un compte rendu de deux fascicules d'un dictionnaire étymologique des langues slaves qui a commencé à paraître sous la direction de Trubačev et qui n'est pas encore terminé. Nous avons estimé alors que ce dictionnaire aurait dû comporter, au total, environ 400 mots présentant un développement phonétique irrégulier dû à la fréquence

(Mańczak 1978), mais jusqu'ici l'évolution d'aucun mot n'y est expliquée de cette manière.

En 2005 a paru un dictionnaire étymologique de la langue polonaise dont l'auteur, Boryś, explique l'évolution de neuf mots par un développement phonétique irrégulier dû à la fréquence. Voici ces mots : *car* « tsar », *chcieć* « vouloir », *człek* « homme », *człowiek* « homme », *ja* « je », *jegomość* « monsieur », *mistrz* « maître », *mość* (par exemple, *wasza królewska mość* « votre majesté »), *niech* « que » (Mańczak 2006).

Le dictionnaire étymologique du polonais de Bańkowski n'est pas encore terminé, mais dans les deux tomes déjà parus l'auteur attribue un développement phonétique irrégulier dû à la fréquence à vingt-deux mots (Mańczak 2001 et 2002).

Ces chiffres ne sont pas impressionnants, pourtant il y a une différence entre ces deux dictionnaires et tous les autres dictionnaires étymologiques. Par exemple dans le *Dictionnaire étymologique de la langue française* de Bloch et Wartburg, dont la 8<sup>e</sup> édition a paru en 1989, l'évolution d'un seul mot, à savoir *que*, est expliquée par la fréquence (« Son fréquent usage l'affaiblit en *que* »), tandis que, dans le *Etymologisches Wörterbuch der deutschen Sprache* de Kluge, dont la 23<sup>e</sup> édition a été publiée en 1995, pas un seul mot n'est expliqué de cette manière.

La situation est paradoxale. Depuis un demi siècle, nous répétons incessamment que, dans toutes les langues du monde, la forme des mots dépend aussi d'un développement phonétique irrégulier dû à la fréquence, mais notre opinion est passée sous silence par la presque totalité des linguistes. Une discussion concernant notre thèse n'a jamais eu lieu : les arguments contre notre conception (s'il y en a) sont tenus en secret au lieu d'être publiés. Tout auteur d'un dictionnaire étymologique devrait demander à un statisticien si les données que nous présentons prouvent, oui ou non, qu'un lien existe entre les réductions irrégulières dans les morphèmes, mots et groupes de mots et la fréquence d'emploi, mais personne ne le fait. Une méfiance irrationnelle envers la statistique freine le progrès de la science du langage.

Witold Mańczak  
Instytut Filologii Romańskiej UJ  
al. Mickiewicza 9/11  
PL – 31-120 Kraków

---

**R é f é r e n c e s**

- Lehnert M. 1971, *Rückläufiges Wörterbuch der englischen Gegenwartssprache*, Leipzig.
- Mańczak W. 1969, *Le développement phonétique des langues romanes et la fréquence*, Kraków.
- 1977, *Słowiańska fonetyka historyczna a frekwencja*, Kraków.
- 1978, c. r. du « Ètimologičeskij slovař slavjanskich jazykov », *Slavia Occidentalis* 35, p. 151-153.
- 1987, *Frequenzbedingter unregelmässiger Lautwandel in den germanischen Sprachen*, Wrocław.
- 2001, Nieregularny rozwój fonetyczny spowodowany frekwencją w słowniku Bańkowskiego, *Prace Językoznawcze Uniwersytetu Śląskiego* 26, p. 153-158.
- 2002, Nieregularny rozwój fonetyczny spowodowany frekwencją w II tomie słownika Bańkowskiego, *Slavia Occidentalis* 59, p. 51-56.
- 2006, c. r. de Boryś, « Słownik etymologiczny języka polskiego », *Język Polski* 86, p. 144-148.
- Thorndike E.L., I. Lorge 1944, *The Teacher's Word Book of 30,000 Words*, New York.